

# Evaluating the Validity of the Dichotomous Choice Question Format in Contingent Valuation

JOHN LOOMIS<sup>1</sup>, THOMAS BROWN<sup>2</sup>, BEATRICE LUCERO<sup>2</sup> and  
GEORGE PETERSON<sup>2</sup>

<sup>1</sup>*Department of Agriculture and Resource Economics, Colorado State University, Fort Collins, CO 80523, USA;* <sup>2</sup>*Rocky Mountain Forest and Range Experiment Station, USDA Forest Service, Fort Collins, CO 80523, USA*

Accepted 1 November 1996

**Abstract.** Hypothetical and actual cash willingness to pay (WTP) for an art print were elicited with dichotomous choice and open-ended question formats. Comparing hypothetical and actual dichotomous choice responses using both a likelihood ratio test and the method of convolutions suggests we reject equality at the 0.05 but not the 0.01 level. Hypothetical WTP was roughly two times actual WTP with the dichotomous choice format. There were no significant differences between the open-ended and dichotomous choice question formats when both were used to estimate hypothetical WTP or both used to estimate actual WTP.

**Key words:** hypothetical bias, willingness to pay, experimental economics

## Introduction

Valuation of public goods frequently relies upon the use of a structured survey technique called the Contingent Valuation Method (CVM). A long-standing criticism of CVM is that stated willingness to pay (WTP) may be a poor indicator of actual WTP (Diamond and Hausman 1994). One possible reason for the alleged poor performance of CVM and a criticism of past CVM studies is the use of an open-ended WTP question format, where respondents are directly asked to state their maximum WTP for the good. This question format is certainly different from the normal price taking behavior where consumers react to posted prices. It is also different from public good decisions voters make in a referendum (Hoehn and Randall 1987). To better mimic price taking in market behavior, it is now common to ask respondents whether they would pay a given dollar amount that varies (randomly) from respondent to respondent and then statistically infer the maximum WTP (Hanemann 1984). This question format is referred to as dichotomous choice (DC) or take-it-or-leave-it.

However, use of the DC format is not without its costs. First, compared to directly asking individuals their WTP, the DC format is statistically inefficient, requiring substantially larger samples for the same level of precision. If in-person interviews are to be used as recommended by the NOAA panel (Arrow et al. 1993),

then the cost savings of open-ended questions could be substantial. Second, the DC question format may allow biases unique to the yes/no format. For example, the DC format might result in symbolic votes in favor of the environmental program, not because the respondent would pay the posited price for it, but rather to register their support for providing the public program (Brown et al. 1996). The DC format may also encourage "yea saying", whereby the posited bid is accepted as a cue of what is a reasonable payment (Kanninen 1995; Mitchell and Carson 1989).

Although there are several advantages and disadvantages of dichotomous choice WTP questions, one decisive factor is likely to be the validity of this question format. Although there are many types of validity (Mitchell and Carson 1989), comparison of actual WTP with stated WTP is one which carries much weight among economists. Comparisons of stated and actual WTP are often called tests of criterion validity, where actual WTP is accepted as the criterion.

Unfortunately, there have been relatively few criterion validity tests of the dichotomous choice method. The first field validity tests of dichotomous choice CVM for estimating WTP were performed for deer hunting permits (Bishop et al. 1992; Welsh 1986). As reported in Bishop et al. (1992), dichotomous choice CVM estimates of mean WTP exceeded actual WTP by a factor of just 1.23 and the two estimates were not found to be significantly different. Kealy et al. (1988) found that undergraduates' hypothetical WTP for a chocolate bar was 1.4 times actual WTP using DC question format and the differences were statistically significant. Champ et al. (1995) found that dichotomous choice estimates of hypothetical WTP for a *public* good exceeded actual WTP by a factor of six. Finally, Cummings et al. (1995) found responses to hypothetical dichotomous choice responses at the one bid amount asked overstated actual buying behavior for juicers, calculators and chocolates. However, Smith (1994) identifies several shortcomings in the Cummings et al. (1995) experiment that may have led to their results. These include failure to clarify to the respondent in the hypothetical treatment not to report what the juicer was worth, but rather whether they would intend to buy one at the price. Further, no distinction was made between respondents who currently owned a juicer and hence were probably not in the market for another one, and respondents who did not own a juicer. The experiments reported here provide wording to respondents in an attempt to overcome these two limitations.

The objectives of this research are to: (a) test whether there is a difference between dichotomous choice responses in a hypothetical market versus a real market for a private good and (b) compare the performance of the dichotomous choice and open-ended WTP question formats in real and hypothetical markets for the same private good.

## 2. Research Design

The laboratory experiment reported here involves four independent treatments:  
No. 1: Dichotomous choice (dc) WTP in the hypothetical (hyp) market.

No. 2: Dichotomous choice WTP in the real cash market.

No. 3: Open-ended (oe) WTP in hypothetical market.

No. 4: Open-ended WTP in real cash market.

The hypotheses to be tested are:

$$H_0: WTP(hyp-dc) = WTP(real-dc) \quad (1)$$

$$H_0: WTP(real-dc) = WTP(real-oe) \quad (2)$$

$$H_0: WTP(hyp-dc) = WTP(hyp-oe) \quad (3)$$

$$H_0: WTP(hyp-oe) = WTP(real-oe) \quad (4)$$

### 3. Statistical Techniques

Maximum WTP is not directly observed in the DC approach but it can be estimated parametrically or calculated non-parametrically. Parametrically, the two most common approaches are Hanemann's (1984) utility difference and Cameron's variation function (1988). McConnell (1990) has shown that these two approaches are equivalent with linear specifications of the random utility model and constant marginal utility of income, and so we adopt Hanemann's as a matter of computational convenience. Hanemann (1984) views CVM respondents using a utility difference approach when they decide whether to answer 'yes' or 'no' at the stated bid amount (\$BID). If the utility difference is logistically distributed, a logit model of the probability of a YES response is related to the respondent's bid amount (\$BID) and attitude/demographic variables ( $Z$ ) as in Equation (5):

$$\log[\text{Prob}(\text{YES})/(1 - \text{Prob}(\text{YES}))] = B_0 - B_1(\$BID) + B_2(Z_1) + \dots + B_n(Z_n). \quad (5)$$

WTP is the area under the cumulative distribution function (CDF or  $g(BID)$ ) between zero and infinity:

$$WTP = \int_0^{\infty} [1 - g(\$BID)] d\$BID \text{ when } WTP > 0. \quad (6)$$

To calculate the mean WTP from the truncated logistic distribution the formula for the mean of a non-negative random variable is used (Hanemann 1989):

$$\text{Mean WTP} = 1/B_1 * (\ln(1 + \exp(B_0 + \Sigma(B_n(Z_n))))) \quad (7)$$

The median is provided by:

$$\text{Median WTP} = (B_0 + \Sigma(B_n(Z_n)))/B_1 \quad (8)$$

where  $B_n$  is the vector of coefficients and  $Z_n$  are the sample means of the associated independent variables.

If the utility difference is assumed to be normally distributed, then the probit model is estimated. The probit model is given by:

$$F^{-1}(\pi_i) = B_0 - B_1(\$BID) + B_2(Z_1) + \dots + B_n(Z_n) \quad (9)$$

where  $F^{-1}(\pi_i)$  is the inverse of the standard normal cumulative distribution function (Kmenta 1986: 553). Median WTP is calculated as in Equation (8).

Mean WTP can also be calculated using a non-parametric approach proposed by Kristrom (1990). Based on the proportion of yes responses at each bid amount, the area under what Kristrom calls the 'empirical survival function' is calculated.

Several approaches are available for testing the equality of hypothetical and actual cash WTP estimated from DC questions administered to two independent samples. First, we compare confidence intervals about estimates of WTP(real-dc) and WTP(hyp-dc). This was done for the means estimated parametrically based on the approach of Park et al. (1991) and estimated non-parametrically, using the approach of Duffield and Patterson (1991). If the confidence intervals do not overlap, we may conclude that hypothetical and actual WTP are different. A second approach involves a test of coefficient equality of the logit equations used to estimate WTP. As shown in Equation (7), WTP depends directly on the estimated coefficients in the logit equation. We test for equality of the logit coefficients using a likelihood ratio (LR) test. The null hypothesis is our LR test is:  $\mathbf{Bn(h)} = \mathbf{Bn(a)}$ , where  $\mathbf{Bn(h)}$  and  $\mathbf{Bn(a)}$  are the coefficients in the logit equations for hypothetical WTP and actual cash WTP, respectively. The test is carried out by comparing the log likelihood of a single logit equation (i.e., the restricted model) estimated by combining or pooling observations from both the actual cash and hypothetical WTP responses to the sum of the log likelihood of the two (cash and hypothetical) logit equations estimated separately (i.e., the unrestricted model). The LR test follows a chi-square distribution with degrees of freedom equal to the sum of the number of coefficients in the two unrestricted models minus the number of coefficients in the restricted model.

A third approach, known as the method of convolutions, employs a formal statistical test of the differences in empirical WTP distributions derived from DC data (Poe et al. 1994). These authors note that their method is less prone to type II error than a comparison of confidence intervals and more relevant to comparisons of mean WTP than the likelihood ratio test of logit coefficient equality. The method of convolutions determines if there is a statistically significant difference between two simulated WTP distributions. According to Poe et al. (1994: 907) it accommodates any distributional form. The method involves calculating the probability of all possible differences (i.e., the convolutions) between discrete values in the two distributions. The method then tests whether the  $1 - \alpha$  confidence interval for this convolution or set of differences includes zero. In addition, the method calculates an alpha level for rejecting the null hypothesis of equality of the two distributions.

A fourth approach determines whether the odds of agreeing to pay the bid amount (e.g.,  $[(\text{Pr(Yes)})/(1 - \text{Pr(Yes)})]$ ) are equal between the hypothetical and

actual dichotomous choice treatments using the Cochran–Mantel–Haenszel test (Snedecor and Cochran 1980: 210–213). This non-parametric contingency table test will be used to test hypothesis no. 1.

## **4. Experimental Design**

### **4.1. PARTICIPANTS**

University clerical and administrative staff in academic and non-academic units were recruited and paid \$20 for attending a 45-minute session held on campus. The four sessions were conducted before work, at lunch, and after work on three consecutive days. The winning price was not announced at the end of the sessions, and sessions involved unrelated departments. There were 32 people in the open-ended real cash WTP session and 33 people in the open-ended hypothetical WTP session. Because DC–CVM only records whether an individual's WTP is greater or less than their bid amount, the DC–CVM treatments required larger samples. Each DC–CVM treatment had a total sample of 56 people, which was split into smaller groups of 28 people each. The same researcher conducted all of the sessions, closely following a script for each session.

### **4.2. SELECTION OF MARKET GOOD**

We sought several desirable characteristics of the good to be used in the experiment. First, the good needed to be one that was infrequently purchased and for which there was a fair amount of price dispersion in the market, so that most people would not be familiar with the market price. The objective was to minimize the likelihood that the respondent would simply try to use the market price in determining whether to answer yes or no. Second, we desired a good that had readily observable characteristics, so there would be minimal ambiguity about the product.

Given these characteristics we chose a signed wildlife art print as our good. Art prints can range in price from a few dollars to several hundred dollars and the full extent of the product is completely observable. From among several selections of wildlife art, a signed print of a wolf standing in the forest was selected based on university staff responses to a short questionnaire.

A combination of three features, when taken collectively, distinguishes this experiment from others. First, adults rather than students are the subjects. Second, a signed wildlife art print as the good. This type of good helps to avoid respondents being contaminated by prior knowledge of market prices because the good is bought only infrequently, sold primarily in specialty stores and naturally has a large range of prices. Furthermore, we have employed multiple statistical tests including the first use of the method of convolutions to compare the distribution of WTP from hypothetical and actual dichotomous choice responses estimated with a logit model.

### 4.3. WORDING OF WTP QUESTIONS

The wording of the DC CVM question in the hypothetical market treatment was: *You are being asked to participate in a hypothetical sealed bid auction for this art print. We would like to know if you would pay the dollar amount in question no. 4 below to take this art print with you at the end of this session, if this one art print were actually for sale.*

*At this time in the survey, we are not asking what you think the art print might sell for in a store or what you think its fair price is. Rather, we want to know whether you would honestly be prepared to pay the dollar amount stated in question no. 4 below right now to buy the art print you are being shown; if you would really be required to pay your bid amount with cash, write a check today, or sign a Promissory Note payable on or before August 19. Please take into consideration your budget and what you can afford to pay. If the price in question no. 4 is different from what you judge a fair price to be, that is OK. We want to know if you would actually be prepared to pay the price listed in question no. 4 for the art print.*

*Take a few moments to think about whether you honestly would be prepared to pay the printed dollar amount for this art print if it were being offered for sale to you today. Although the question is hypothetical, we want you to answer as if it were for real – as if you were participating in a real sealed-bid auction and would really be required to pay the printed dollar amount. If only one person answers YES, he or she would have obtained the print at the stated price on the survey. If there is more than one person stating YES we will have additional questions to determine who would have been the highest bidder.*

*4. Would you really be prepared to pay \$BID for this art print?*

*\_\_\_ YES, I would pay this amount, \_\_\_ NO, I would not pay this amount.*

The prelude to the WTP question is different from those of most past CVM questions (particularly those dealing with market goods) in that we asked respondents *not* to simply estimate what they think the good sells for and to act as if the commitment to pay was real. These two statements were included after debriefing sessions following pretests revealed that respondents were using different criteria to answer the hypothetical as opposed to the real cash WTP questions. See Appendix A for more discussion of this pretesting procedure.

The wording of the open-ended WTP question in the hypothetical treatment was the same as the dichotomous choice in terms of the introductory lead in. The key differences related to question format, such as requesting they state their maximum WTP and informing them that the highest bidder would have received the print at the price he or she stated. This open-ended WTP question format is similar to what is used in many CVM surveys and resembles a first price auction with private values (Davis and Holt 1993). We recognize that the open-ended WTP question format is not incentive compatible. However, we also use this format to elicit open-ended WTP in the actual cash market. Therefore we expect any *difference* in these two

open-ended WTPs to be due to the hypothetical nature of the question, because the lack of incentive compatibility is present in both the actual cash and hypothetical open-ended WTP questions. Support for this belief can be seen in the results of Neill et al. (1994), where the incentive compatible Vickery auction format gave hypothetical WTP nearly identical to first price open-ended WTP responses.

The wording of the real market DC CVM question used in the independent actual cash session was: *We are now going to conduct a real auction. If you wish to actually buy the art print at the price stated below, answer YES in question no. 4. If you are the only person who answers YES, you will be required to buy the art print at the stated price. If there is more than one person stating YES, we will have additional questions to determine the highest bidder. We will accept cash or check for your purchase. We understand that you may not have anticipated the need to bring cash or your checkbook with your today, so we will also accept a signed Promissory Note payable on or before August 19.*

*In any case, the successful buyer will be able to take the art print with them at the end of this session. Now take a few moments to think about what having this art print would be worth to you. If you want to buy the art print at the stated price on the sheet, answer YES. If you don't want to purchase the art print at this price, answer NO.*

4. Are you prepared to pay \$*BID* for this art print?

\_\_\_ YES, I will pay this amount. \_\_\_ NO, I will not pay this amount.

In both the hypothetical market and the real market, each person's answer sheet contained one of ten different prices ranging from \$2 to \$120, but centered around the mean of the pre-test open-ended WTP responses, \$38.

The wording of the open-ended WTP question in the real market was: *We are now going to conduct a real auction. This art print will be sold to the highest bidder here today. Only one of these prints will be sold at this auction.*

*After all bids have been collected, the person who is the highest bidder will be announced and he or she will be obligated to purchase the print at his or her bid price. We will accept cash or check for your purchase. We understand that you may not have anticipated the need to bring cash or your checkbook with you today, so we will also accept a signed Promissory Note payable on or before August 19. In any case, the highest bidder will be required to pay his or her bid amount and will then be able to take the art print with him or her at the end of this session. What is the most you are prepared to pay for this art print? I bid \$\_\_\_.*

#### 4.4. SETTING OF THE EXPERIMENT

All the sessions were held in a classroom with participants sitting at every other seat to maintain privacy and eliminate any discussion with each other. At the beginning of the session, individuals were shown the art print and then asked to rate, using a five-point Likert scale (ranging from strongly disagree to strongly

agree) whether they liked the print (variable LIKE) and whether they were in the market for an art print such as this (variable MARKET). Next, individuals were instructed to read and complete the WTP question. When everyone had finished, the response sheets were passed forward. The respondents then filled out a sheet on their demographics. In the real markets, the person who was the highest bidder was announced and asked to come forward to complete his or her purchase in front of the group, but the winning price was not announced. Individuals were allowed to pay with cash or check or sign a promissory note payable within three weeks.

## 5. Results

### 5.1. COMPARISON OF DEMOGRAPHICS ACROSS SESSIONS

Before comparing estimates of WTP, we tested whether the four samples were significantly different from each other in terms of standard demographics. One-way ANOVAs were performed for education ( $F = 1.81$ ,  $p = 0.15$ ), age ( $F = 0.92$ ,  $p = 0.43$ ) and income ( $F = 1.28$ ,  $p = 0.282$ ). As indicated by the  $p$  values, the samples are not significantly different.

### 5.2. ESTIMATED EQUATIONS

In order to calculate WTP and perform the statistical comparisons of hypothetical and real market behavior, logit equations were estimated for the two dichotomous choice treatments. We hypothesized that WTP for the art print was positively related to how strongly respondents agreed with the statement that they were in the market for this type of art print (MARKET). This variable, which had response categories ranging from 1–5, had a mean of 3.2 and 2.9 in the hypothetical and actual treatments, respectively. How strongly they liked the art print (LIKE) also had response categories ranging from 1–5. Income (INC) measured in thousands and AGE of the respondents were included as demographic variables. Equations (10) and (11) provide the logit equations for hypothetical and real markets:

$$\text{YPAY(hyp)} = -10.77 - 0.2578(\text{\$BID}) + 1.96(\text{MARKET}) + 7.84(\text{LIKE}) - 0.537(\text{AGE}) + 0.09(\text{INC}) \quad (10)$$

(*t* statistics)    (−1.75) (−2.38)                    (1.85)                    (2.27)                    (−2.12)                    (1.55)

This logit equation's goodness of fit statistic, the chi-square, equals 56.6, which is significant at the 0.01 level.

$$\text{YPAY(real)} = -7.92 - 0.1787(\text{\$BID}) + 1.44(\text{MARKET}) + 1.37(\text{LIKE}) - 0.04(\text{AGE}) + 0.05(\text{INC}) \quad (11)$$

(*t* statistics)    (2.05) (−2.56)                    (2.47)                    (1.88)                    (−0.88)                    (1.36)

This logit equation's chi-square equals 36.6, which is significant at the 0.01 level.

Equations (12) and (13) provide the probit equations for the hypothetical and real markets.



$$\text{YPAY(hyp)} = -6.11 - 0.1499(\text{\$BID}) + 1.15(\text{MARKET}) + 4.51(\text{LIKE}) - 0.31(\text{AGE}) + 0.05(\text{INC}) \quad (12)$$

(t statistics)    (-1.75) (-2.42)            (1.83)            (2.28)            (-2.10)            (1.49)

$$\text{YPAY(real)} = -4.11 - 0.0809(\text{\$BID}) + 0.745(\text{MARKET}) + 0.678(\text{LIKE}) - 0.02(\text{AGE}) + 0.02(\text{INC}) \quad (13)$$

(t statistics)    (2.05) (-3.12)            (2.57)            (1.82)            (-0.68)            (1.16)

In both the logit and probit models, the \$BID variable is significant and negatively related to the probability of a 'yes' response in both hypothetical and actual markets, whereas being in the market and liking the good increased the probability of a 'yes' response. The pattern of variable significance is identical between the logit and probit models.

While it is not strictly necessary to estimate WTP equations when an open-ended question format is used, comparability with WTP estimated from the logit equation may be improved by using an equivalent behavior model. Equations (14) and (15) present WTP equations estimated using an equivalent specification to the logit models:

$$\text{WTP(hyp-oe)} = -72.1 + 3.32(\text{MARKET}) + 12.9(\text{LIKE}) + 1.1(\text{AGE}) - 0.094(\text{INC}) \quad (14)$$

(t statistics)    (3.58)    (1.04)            (2.87)            (2.90)            (-0.67)

$$R^2 = 0.59 \quad F = 9.93$$

$$\text{WTP(real-oe)} = -3.09 + 3.71(\text{MARKET}) + 1.36(\text{LIKE}) - 0.022(\text{AGE}) + 0.064(\text{INC}) \quad (15)$$

(t statistics)    (-0.18)    (1.41)            (0.36)            (-0.08)            (0.82)

$$R^2 = 0.21 \quad F = 1.76$$

The equation explaining the hypothetical WTP responses performs reasonably well, while the equation explaining the actual cash open-ended WTP responses is far less satisfactory.

Table I summarizes the means and 95% confidence intervals (CI) for all four treatments. The parametric results represent estimates calculated from the logit and OLS regression for dichotomous choice and open-ended responses, respectively. The non-parametric results represent Kristrom's estimator for the dichotomous choice and the average of the raw data for the open-ended responses. Estimates of WTP from the probit models are not shown as they are not statistically different from the logit estimates.

The results indicate that the parametric DC estimate of hypothetical WTP is 2.5 times the DC estimate of actual WTP and two times the open-ended estimate of actual WTP. The non-parametric DC estimate of hypothetical WTP is 3 times the estimate of actual WTP and 2.3 times the open-ended estimate of actual WTP. A similar relationship is evident from examining the differences in median WTP between treatments, in spite of the rather direct wording in the hypothetical WTP treatments to consider their budget and act as if the auction were real.

Table I. Comparison of hypothetical and actual WTP.

Treatment	Estimator	Sample size	Mean (median) WTP by treatment (95% CI)			
			Dichotomous choice		Open-ended	
			Hypothetical	Real market	Hypothetical	Real market
1.	Logit	52	28(28) (20–37)			
	Non-parametric		33 (20–46)			
2.	Logit	55	11(9) (6–22)			
	Non-parametric		11 (9–14)			
3.	OLS	33			26(20) (19–33)	
	Non-parametric				26(20) (17–35)	
4.	OLS	32				14(14) (12–16)
	Non parametric					14(10) (10–18)

### 5.3. RESULTS OF HYPOTHESIS TESTS

#### 5.3.1. $H_0: WTP(hyp-dc) = WTP(real-dc)$

As can be seen in Table I, the 95% confidence intervals (CIs) for hypothetical and actual WTP derived from the logit estimates of the DC responses overlap in the tails, making this test inconclusive. A LR test of the equality of the logit coefficient yields a chi-square of 16.11, to be compared with critical chi-square values of 12.59 at the 0.05 significance level and 16.812 at the 0.01 level. Thus we reject the null hypothesis that the independent variables affect the dependent variable of Equation (10) in the same manner that they affect the dependent variable of Equation (11) at the 0.05 significance level, but we cannot reject equality at the 0.01 level. The method of convolutions for comparing WTP distributions estimated from hypothetical and real markets indicates we should reject equality at the 0.05 level but not at the 0.01 level. This is consistent with the LR test. As shown in Table II, the Mann–Whitney test suggests the medians are statistically different at the 0.01 level. The Cochran–Mantel–Haenszel non-parametric, contingency table test also rejects hypothesis no. 1 as the odds of a respondent providing a Yes response is significantly ( $P = 0.018$ ) greater in the hypothetical treatment than in the cash treatment.

Table II. Probability levels associated with testing the equality of actual and hypothetical WTP.

Hypothesis	Mann-Whitney test of medians	Method of convolution test	Coefficient equality test
WTP(hyp-dc) = WTP(real-dc)	$P = 0.006$	$P = 0.05$	$P = 0.05$
WTP(real-dc) = WTP(real-oe)	$P = 0.77$	NA	NA
WTP(hyp-dc) = WTP(hyp-oe)	$P = 0.48$	NA	NA
WTP(hyp-oe) = WTP(real-oe); (Est WTP)	$P = 0.04$	NA	$P = 0.01$
WTP(hyp-oe) = WTP(real-oe); (raw data)	$P = 0.09$	NA	NA

### 5.3.2. $H_0: WTP(real-dc) = WTP(real-oe)$

The CI for WTP(real-oe) from either the parametric or non-parametric (e.g., mean of the raw data) are completely contained within the CI for WTP(dc-real) using the parametric approach, suggesting that actual WTP estimated using the two question formats are not significantly different. The Mann-Whitney test of the difference in the medians is also insignificant ( $P = 0.77$ ). Thus both of these approaches (CIs and Mann-Whitney) indicate no difference between actual cash WTP elicited with open-ended and actual cash estimated from dichotomous choice.

### 5.3.3. $H_0: WTP(hyp-dc) = WTP(hyp-oe)$

The Mann-Whitney test of the medians yields a  $P$  value of 0.48, indicating no significant difference in the medians between hypothetical WTP estimated using dichotomous choice and open-ended question formats. The mean and CIs of hypothetical WTP elicited from dichotomous choice logit model are nearly identical to the hypothetical WTP from the open-ended question format (\$28 vs. \$26 for means of hyp-dc and hyp-oe, respectively, and \$20–37 vs. \$19–33 for 95% CIs of hyp-dc and hyp-oe, respectively).

### 5.3.4. $H_0: WTP(hyp-oe) = WTP(real-oe)$

Although not the main focus of our paper, Table II shows that median WTP elicited in the two open-ended treatments are statistically different at the 0.09 level using the Mann-Whitney test on the raw data. The means are statistically different at the 0.024 level using a  $t$ -test on the raw data. Using a Chow test of coefficient equality of the regressions in Equations (14) and (15), equality of hypothetical and actual WTP elicited using the open-ended question format is rejected at the 0.01 level ( $F = 4.97$ , while the critical  $F = 3.34$ ).

## 6. Source of Differences Between Hypothetical and Actual WTP

It is interesting to hypothesize about the source of the differences between hypothetical and actual WTP. Is this a constant difference found across the entire distribution

or a pronounced divergence primarily in the right-hand tail? To investigate this, we scrutinized the largest hypothetical open-ended WTP responses. Our deletion criterion was WTP responses over \$56. Four hypothetical bids (about 10% of the sample) were in excess of \$56. Figure 1 illustrates the fairly close correspondence of hypothetical and actual WTP up to the cut-off, and the excess untrimmed hypothetical responses at the high bid amounts. The effect of trimming these four bids is that mean (median) hypothetical WTP drops from \$26 (\$20) to \$19 (\$15). The result is that both the mean and the median hypothetical open-ended WTP are no longer statistically different from corresponding actual values, with  $P$  values of 0.27 for the mean and 0.356 for the median. Much of the difference appears to be in the right tail, representing the responses of a small minority of respondents. As suggested by a reviewer, we investigated the demographics of these four respondents versus the rest of the sample to see if there were any obvious demographic differences. These individuals were lower income (their mean income was \$19 000 less than the sample average), had below average education (two of the four had just 12 years of education) and were older than average (their age was 10 years older than sample average). The lower income makes it surprising these people would be high bidders. We hypothesize their high hypothetical bids may be due to less appreciation of the need to formulate an accurate measure of the WTP in an experimental setting or perhaps less ability to formulate their WTP when the payment is hypothetical. Future experiments might attempt debriefing sessions with respondents that formulate bids substantially higher than other members of the sample to more thoroughly investigate the source of such high bids.

## 7. Discussion and Conclusion

This study can be compared with those of Bishop et al. (1992) and Brown et al. (1996), as they also obtained dichotomous choice and open-ended estimates of WTP in hypothetical and actual markets, and also used independent samples for each of the four experimental conditions. The principal difference among the studies was that ours and Bishop et al.'s valued a private good whereas Brown et al.'s valued a public good. Perhaps most notably, all three studies found that *actual* WTP was roughly the same regardless of whether the open-ended or dichotomous choice format was used (Bishop et al. do not report a statistical test, but the other two studies found no significant difference). With respect to hypothetical WTP the evidence is more mixed, with our study and Bishop et al. showing no difference but other private good (e.g., recreation) studies and public good studies showing the dichotomous choice estimate exceed the open-ended estimate (see Brown et al. (1996) for a list of the studies).

The statistical tests also suggest that hypothetical WTP estimated using the dichotomous choice method exceeds actual cash WTP for the art print used in our experiment. The ratio of hypothetical to actual WTP using dichotomous choice was 2.54 to one; this difference is statistically significant at the 0.05 level but not

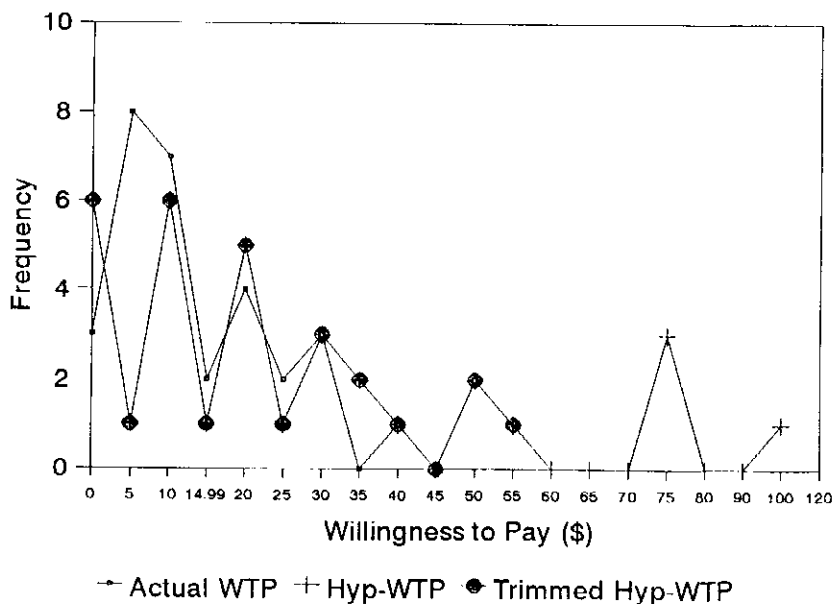


Figure 1. Hypothetical (crosses) + and actual open-ended WTP (small squares) □ with and without trimming.

the 0.01. This ratio is higher than for the open-ended question format, which was 1.85 to one. Our overall results are fairly consistent with those of Kealy et al. (1993), and Brown et al. (1996) in that hypothetical WTP is statistically greater than actual cash WTP whether the open-ended or dichotomous choice format is used. Taken together, these studies suggests that it may be premature to abandon use of open-ended WTP questions. Clearly additional replications with different goods and larger samples are desirable before one may determine whether dichotomous choice question formats should always be preferred to open-ended questions in CVM as recommended by the NOAA panel.

### Acknowledgements

This research was supported under the W-133 Regional Research Project. We would like to thank Helen Neill, University of Nevada–Las Vegas, and Philip Ganderton, University of New Mexico, for their valuable suggestions on an earlier draft of our experimental design. Rudy King and Paul Mielke provided expert advice regarding statistical tests. We would especially like to thank Michael Welsh of HBRIS Inc., for providing the software routines and assistance to allow us to perform the convolutions statistical test on the dichotomous choice data. Patricia Champ provided numerous suggestions that clarified the manuscript. These helpful individuals are in no way responsible for what follows. Constructive suggestions

for investigation the robustness of our conclusions to different specifications were provided by an anonymous reviewer.

## Appendix A. Development of experimental procedures

Several pre-test sessions were conducted with university staff to fine-tune the procedures and to better understand the thought processes used in both the actual cash and hypothetical market scenarios. The first two sessions were separate hypothetical and actual cash sessions. Reviewing the responses after these first two sessions indicated a significant disparity between actual and hypothetical WTP. In the next sessions, we first had subjects respond to the hypothetical question. These sheets were collected, demographic sheets filled out and then a new real cash transaction sheet was handed out. Subjects were told they were now participating in a real cash bid situation where they could actually buy the print. If they were the highest yes response (in the dichotomous choice) they were obligated to buy the print at their price. We performed debriefing afterwards to ask respondents why some of them gave substantially different responses from the hypothetical to the actual cash. One key difference was timing of payment. In the actual cash they originally had to pay within a few days, where in the hypothetical no payment deadline had been specified. Respondents told us that their near term cash constraint was often the factor changing their responses. The hypothetical response was closer to what they would pay, after payday or if they did have the cash in their wallet or checking account. The actual cash is what they could afford in the next few days, which did not involve a payday. To put the timing of payment on the same footing, the revised text specified August 19, which involved three weeks including a payday (as staff get paid every two weeks). The August 19 date was used in both the hypothetical and actual cash experiments.

Another area of fine tuning was to establish a seating arrangement and provide verbal instructions not to talk to one another. This eliminated any discussions or sharing of information between participants. This was critical in the dichotomous choice treatment as different respondents have different values. For them to take the bid price as a non-arbitrary fixed price, it was important they did not know that others had different prices. Individuals with questions were asked to hold any non-essential questions until after the session was completed. Subjects with a critical question were asked to raise their hand so that the investigator could quietly address that specific individual's question.

## References

- Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman (1993), 'Report of the NOAA Panel on Contingent Valuation', *Federal Register* 58(10), 4602-4614.
- Bishop, R., M. Welsh, and T. Heberlein (1992), *Some Experimental Evidence on the Validity of Contingent Valuation*. Department of Agricultural Economics, University of Wisconsin.
- Brown, T., P. Champ, R. Bishop, and D. McCollum (1996), 'Which Response Format Reveals the Truth About Donations to a Public Good', *Land Economics* 72(2), 152-166.
- Champ, P., R. Bishop, T. Brown, and D. McCollum (1995), 'Some Evidence Concerning the Validity of Contingent Valuation: Preliminary Results of an Experiment, in W-133, Benefits and Costs Transfer in Natural Resources Planning', *Seventh Interim Report*, compiled by R. Ready, Department of Agricultural Economics, University of Kentucky, Lexington, KY.
- Cameron, T. (1988), 'A New Paradigm for Valuing Nonmarket Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression', *J. Environmental Economics and Management* 15, 355-379.

- Cummings, R., G. Harrison, and E. E. Rutstrom (1995), 'Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?', *American Economic Review* **85**, 260-266.
- Davis, D. and C. Holt (1993), *Experimental Economics*. Princeton University Press.
- Department of Interior (1994), 'Natural Resource Damage Assessments: Proposed Rule', *Federal Register* 43CFR (May 4) **59**(85), 23098-23111.
- Diamond, P. and J. Hausman (1994), 'Contingent Valuation: Is Some Number Better than No Number?', *J. Economic Perspectives* **8**(4), 45-64.
- Duffield, J. and D. Patterson (1991), 'Inference and Optimal Design for a Welfare Measure in Dichotomous Choice Contingent Valuation', *Land Economics* **67**(2), 225-239.
- Hanemann, M. (1984), 'Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses', *American J. Agricultural Economics* **66**(3), 332-341.
- Hanemann, M. (1989), 'Welfare Evaluations in Contingent Valuation Experiments with Discrete Response Data: Reply', *American J. Agricultural Economics* **71**(4), 1057-1061.
- Hoehn, J. and A. Randall (1987), 'A Satisfactory Benefit-Cost Indicator for Contingent Valuation', *J. of Environmental Economics and Management* **14**, 226-247.
- Kanninen, B. (1995), 'Bias in Discrete Response Contingent Valuation', *J. Environmental Economics and Management* **28**(1), 114-125.
- Kealy, M. J., J. Dovidio, and M. Rockel (1988), 'Accuracy in Valuation Is a Matter of Degree', *Land Economics* **64**(2), 158-171.
- Kealy, M. J. and R. W. Turner (1993), 'A Test of the Equality of Close-Ended and Open-Ended Contingent Valuations', *American J. Agricultural Economics* **75**, 321-331.
- Kmenta, J. (1986), *Elements of Econometrics*, Second Edition. New York, NY: Macmillan.
- Kristrom, B. (1990), 'A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Choice Response Valuation Studies', *Land Economics* **66**(2), 135-139.
- McConnell, K. (1990), 'Models for Referendum Data: The Structure of Discrete Choice Models for Contingent Valuation', *J. Environmental Economics and Management* **18**(1), 19-34.
- Mitchell, R. and R. Carson (1989), 'Using Surveys to Value Public Goods: The Contingent Valuation Method', *Resources for the Future*. Washington, DC.
- Neill, H., R. Cummings, P. Ganderton, G. Harrison, and T. McGuckin (1994), 'Hypothetical Surveys and Real Economic Commitments', *Land Economics* **70**(2), 145-154.
- National Oceanic and Atmospheric Administration (1994), 'Natural Resource Damage Assessments: Proposed Rule', *Federal Register*, 59CFR 1062 (January 7).
- Park, T., J. Loomis, and M. Creel (1991), 'Confidence Intervals for the Evaluating Benefit Estimates from Dichotomous Choices Contingent Valuation Studies', *Land Economics* **67**(1), 64-73.
- Poe, G., E. Severance-Lossin and M. Welsh (1994), 'Measuring the Differences in Simulated Distributions: A Convolutions Approach', *American J. Agricultural Economics* **76**(4), 904-915.
- Smith, V. K. (1994), 'Lightening Rods, Dart Boards, and Contingent Valuation', *Natural Resources Journal* **34**(1), 121-152.
- Snedecor, G. and W. Cochran (1980), *Statistical Methods* (7th edition) (Ames, IA: Iowa State University Press).
- Welsh, M. (1986), 'Exploring the Accuracy of the Contingent Valuation Method: Comparison with Simulated Markets', PhD. Dissertation, University of Wisconsin, Madison.